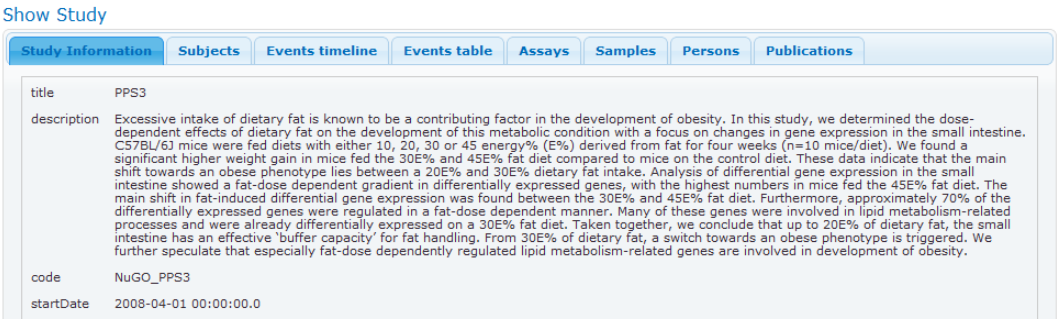
Version: 1.5
Date: 02/03/2012
Written by: Janneke van der Ploeg

GSCF User Guide

Introduction - GSCF

GSCF (Generic Study Capture Framework) is a web application that is designed to store the design of biological studies in a database. GSCF serves as a linking hub for other modules: it links to the data in the various assay modules.

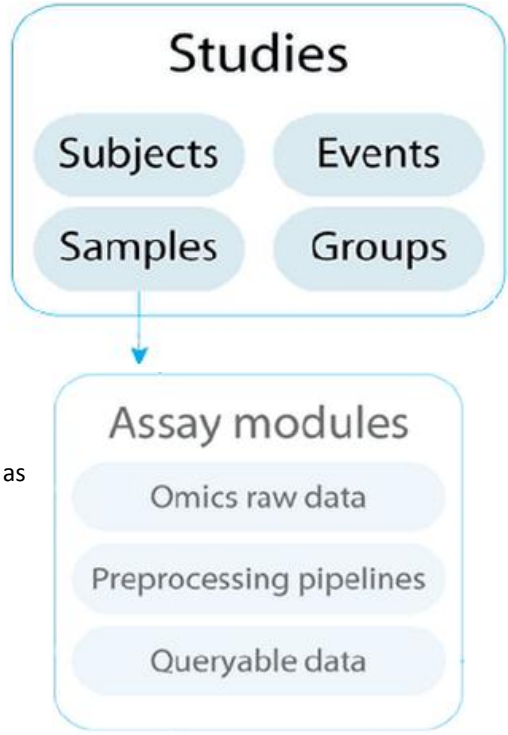
So, before data in modules can be entered, the study design needs to be in GSCF. The following image (headers) illustrates what type of information can be stored in GSCF.



Introduction – assay modules

Users can create samples by their study design in GSCF, and indicate where the assay data is stored. In these assay modules users can easily import, for example bulks of numerical data, from Excel files. The following modules are currently available for GSCF:

- Metagenomics
- Clean transcriptomics (CTD)
- Metabolomics
- Simple Assay Module (SAM)
- Transcriptomics



For example, the Simple Assay Module (SAM) is used to store clinical data, such as Rules Based Medicine assays, but also physical measurements such as body weight of mice. The image below illustrates how GSCF and assay modules are related.

Content

Tips and tricks 3

1. Getting started – simple import and wizard 4

1.1 Create a new study (wizard) 5

1.2 Edit a study 9

1.3 Import part of study design 10

2. Advanced search 14

3. Visualize 16

3.1 How to get GSCF data visualized? 17

3.2 How to get module data visualized? 18

4. Export data 19

4.1 Export GSCF data 19

4.2 Export Assay data 19

5. Access control 21

5.1 Admin and users 21

5.2 Readers and writers 23

6. Templates 26

FAQs 28

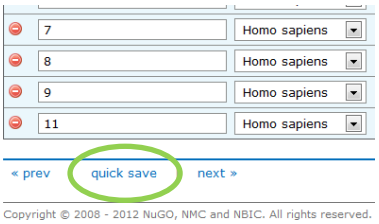
Tips and tricks

General

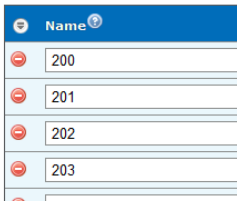
- ✓ Use 'import study data' to import large datasets (for example: many subjects) from an excel sheet into your study (see also Chapter 1.3 Import part of study design).
- ✓ When you want to use Advanced search, and choose *sample* as the output format, note that this search can take several minutes, since the number of samples could be very high.

Study wizard

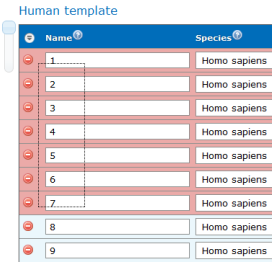
- ✓ Use next and previous (at the bottom of the page) to navigate, or use 'quick save' at any time during the wizard, to complete it at another point in time.



- ✓ Hovering over question marks will show additional information on required fields:
- ✓ Use the upper left icon to select all rows in a table:



- ✓ Select multiple rows by click and dragging over multiple rows:



1. Getting started – simple import and wizard

Using the menu on top of the application, you can either create, view or import studies and study data.

Generally, there are two ways to get your data into the database.

1. Start the study *wizard*, which will guide you step-by-step. This will allow you to describe the study design, all events and sampling events that take place, and on what time. Also into what event groups your subjects are organized. Based on subjects, event groups and events, GSCF will create samples. Samples are the links between GSCF and the assay modules, where you can store specific data (such as metabolomics or clinical chemistry values). (Go to chapter 1.1 for a more detailed description.)

Advantage: full description of your study design, and easy to compare with other GSCF-created studies that are fully described.
Disadvantage: because a lot of information can be stored to describe your study, it takes some time before you can start using the modules. Especially when your study contains a lot of samples (>1000), the wizard can become a bit slow.

2. a. If you have a sample list of all samples in your study, you can use the *short way* to get data into the system. Since samples/samplenames are the link between GSCF and the assay modules, you can simply create a study, and describe study metadata.

b. Then use the importer (Import > A part of study design) to import your Excel sheet, and thereby import your own sample names. These names can then be used in the various assay modules, to add metabolomics data. (Go to chapter 1.3 for a more detailed description.)

Advantage: you can quickly enter data into the module, since only minimal necessary information is imported.
Disadvantage: comparing two studies is more difficult if users only fill in minimal study (meta)data.

➔ Use the content list on page 2, to navigate through this guide
➔ For any questions that could not be answered through this guide, please email support@thehyve.nl

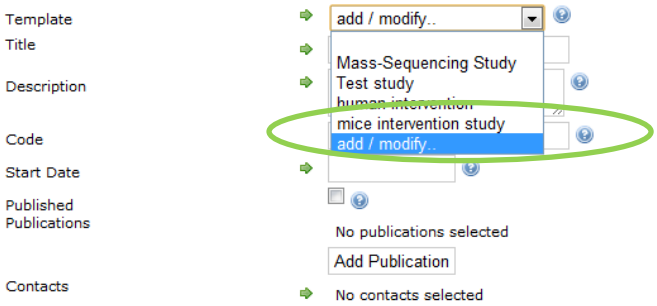
1.1 Create a new study (wizard)

For creating a new study, you have to have a login. See chapter 7 (Access control) on how to become a user. Click on Create > Create a new study using the menu on top. This will start the 6-step study wizard.

Step 1: Start

Define basic properties of your study. Fields indicated by an arrow are required.

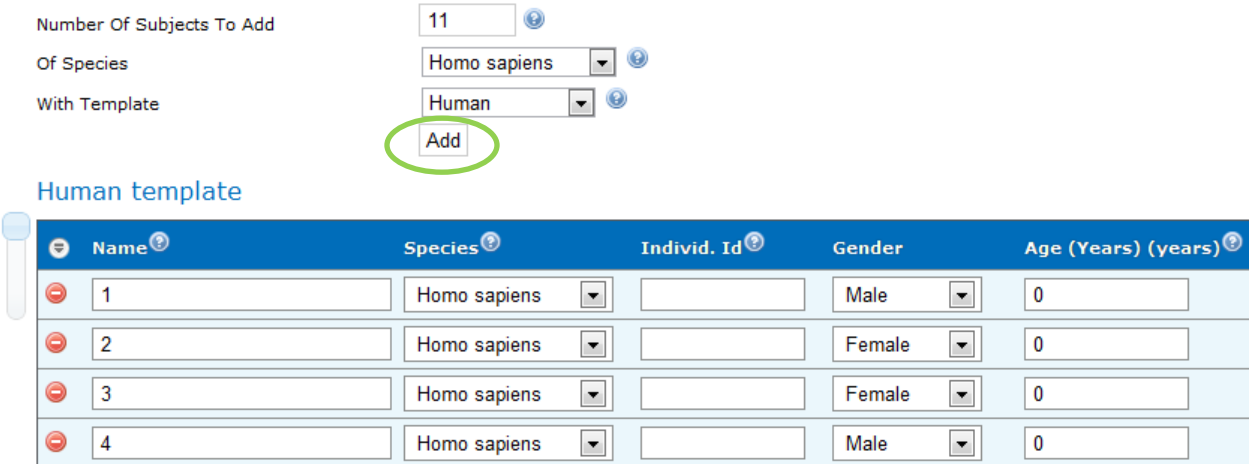
You can choose an appropriate template, or create a new template. This will allow you to add and modify fields.



In some browsers (Firefox 4, Safari and Chrome) fields can be enlarged by dragging the indicated (lower right) corner of the text field.

Step 2: Subjects

Use this step to manually add a number of subjects: report the number of subjects in your study and choose the template. Then add these subjects to your study and fill out the required information in the table.

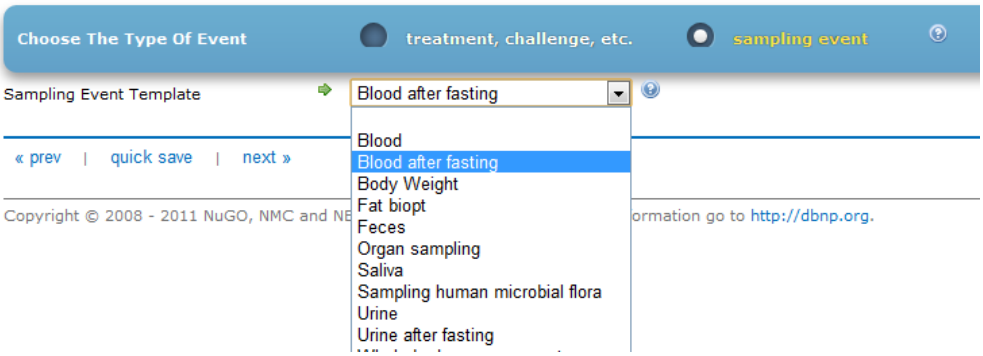


If you have an excel file with subject list, a separate importer can be used. Especially, when you have detailed subject information, this importer can be used very efficiently. See Topic 3 - Import part of study design (page 10).

Step 3: (Sampling) Events

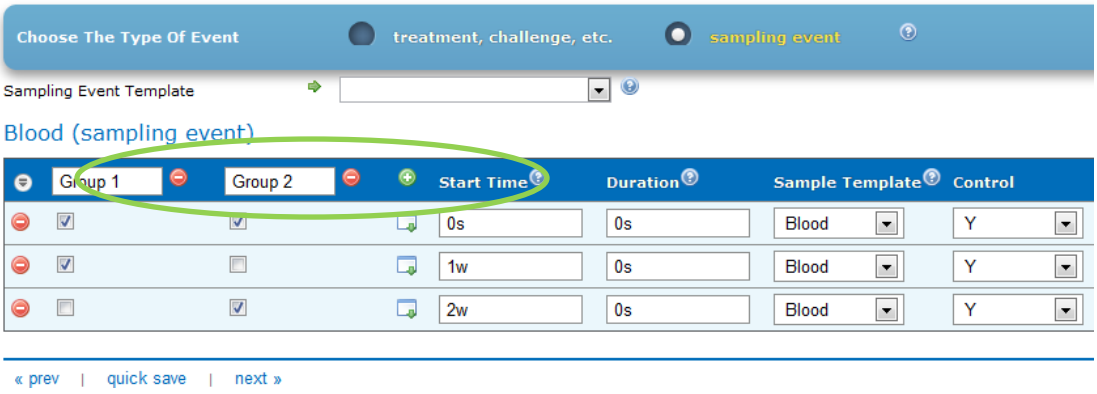
Any treatment, challenge or occurrence in your study can be defined in the events step. An *event* is defined as anything that happens to subject(s) in your study that is not resulting in a sample. Whereas *sampling* events if the event results in extraction of samples, where you want to store measurements upon. First choose between treatment or sampling event:

Then choose the appropriate (sampling) event template:



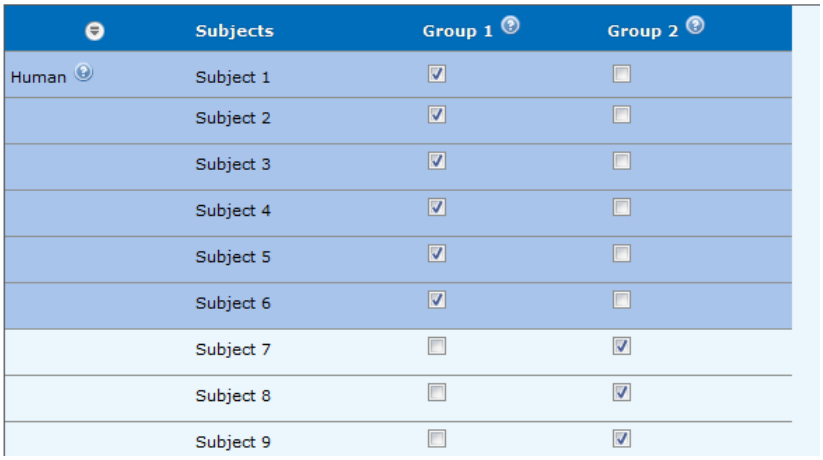
After you have added an event or sampling event, it will appear in a list below, and you can modify this event if necessary. Here, you also need to create any event *groups* that occur in your study. Make sure that each event is linked to a group. Only then subjects can be added into this group (see next step).

- Important: each event should be added separately.
- Important: a combination of sampling events, subjects and groups, will automatically create samples.



Subjects into event groups

Click next: this in-between-step allows you to select subjects and indicate to which event group they belong.



	Subjects	Group 1	Group 2
Human	Subject 1	☑	☐
	Subject 2	☑	☐
	Subject 3	☑	☐
	Subject 4	☑	☐
	Subject 5	☑	☐
	Subject 6	☑	☐
	Subject 7	☐	☑
	Subject 8	☐	☑
	Subject 9	☐	☑

- Note: event groups are created in Step 3. You can fill these groups with subjects between Step 3 and 4.

Step 4: Samples

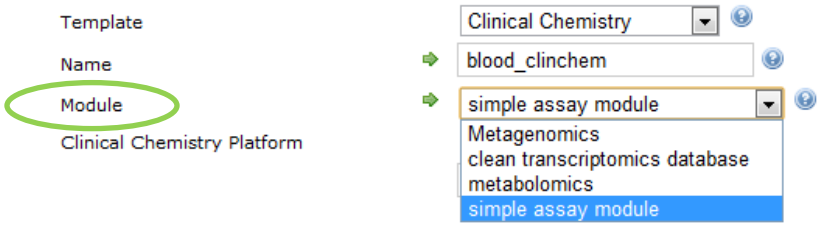
This step gives a table with all the samples that were created. Note that a combination of sampling events, subjects and groups, automatically created your samples.

You can make any adjustments, or fill out empty fields when necessary.

Step5: Assays

In this step assays can be created. Depending on the assay type and module you choose, later on you can attach measurements or data files to a set of samples.

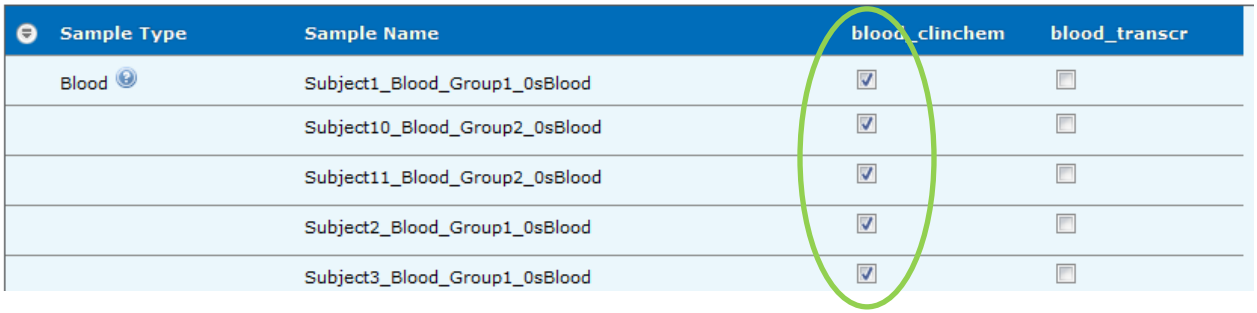
Make sure to select the template (e.g. clinical chemistry), fill in an assay name and select the module (e.g. simple assay module) you wish to use.



Samples into Assays

After you click 'next', an in-between step allows you to select which assays were performed on which samples. Select multiple rows by click and dragging. Use checkboxes to indicate to what assay these samples belong.

(It is possible to attach multiple assay to one sample.)



Sample Type	Sample Name	blood_clinchem	blood_transcr
Blood	Subject1_Blood_Group1_0sBlood	☑	☐
	Subject10_Blood_Group2_0sBlood	☑	☐
	Subject11_Blood_Group2_0sBlood	☑	☐
	Subject2_Blood_Group1_0sBlood	☑	☐
	Subject3_Blood_Group1_0sBlood	☑	☐

Step 6: Confirmation

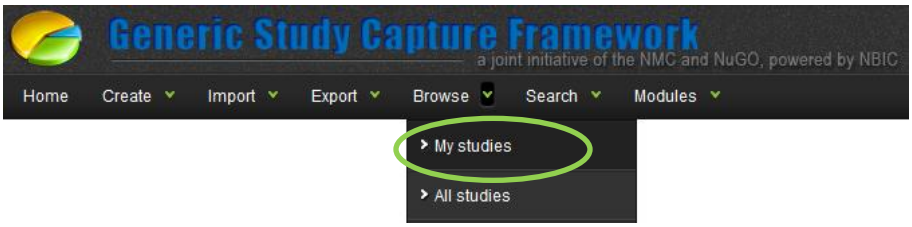
Check your data. Any corrections can be made by going back to the corresponding steps in the wizard.

Step 7. Done

You can view or edit the newly created study, or create a new study.

1.2 Edit a study

To view or edit an already existing study go to Browse > My studies.



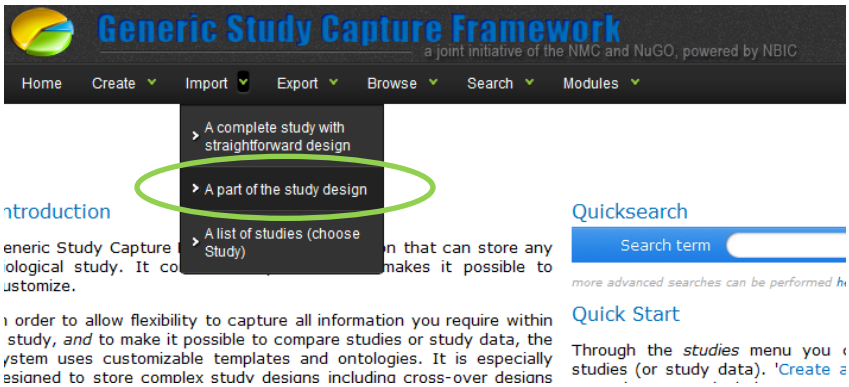
Now you see a list of studies that you have created, or studies where you have the role of reader. Click on the magnifying glass on the left to view a certain study, or click on the pencil in the middle to enter the wizard. This will allow you to edit this study.

1.3 Import part of study design

Instead of walking through the wizard, GSCF also allows users to import a part of the study design. This can be a list of subjects, samples or events.

This is an example of importing subjects:

Click on Import > A part of the study design using the menu on top.



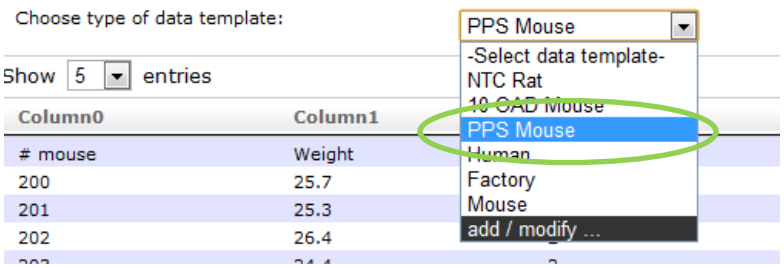
Step 1: Import file

Select your excel file. The excel sheet should have this format: every study, subject, sample or event has its own row.

- It does not matter in *what column* which information is stored, as you can map the field in the template on the columns in your excel sheet.
- It does not matter *how many headers* are included in your excel sheet, as you can define in which row the first subject/ sample or event can be found.

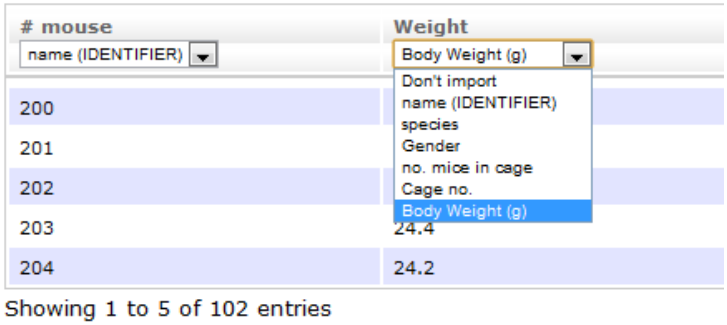
	A	B	C	D	E
1	# mouse	Weight	Cage number	Number of mice in cage	
2	200	25.7	1	2	
3	201	25.3	1	2	
4	202	26.4	2	2	
5	203	24.4	2	2	
6	204	24.2	3	3	

Make sure to select the correct data template; e.g. a mouse subject template for a mouse study. The importer will show a preview of your uploaded excel file.



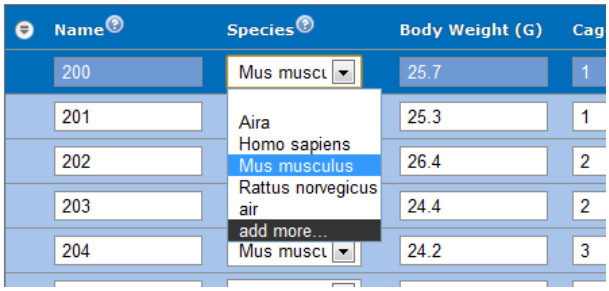
Step 2: Assign properties

Map the information in your excel sheet on the columns, using the dropdowns. If a certain field is missing, go to back to Step 1, where you can modify or add fields for a certain template.



Step 3: Add information

Add information to the table when necessary. Multiple rows can be selected by click and dragging.



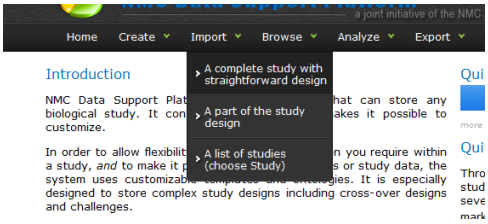
Step 4: Confirmation

Check whether the information is correctly mapped and click > Next. Click on the study title to view your imported study data.

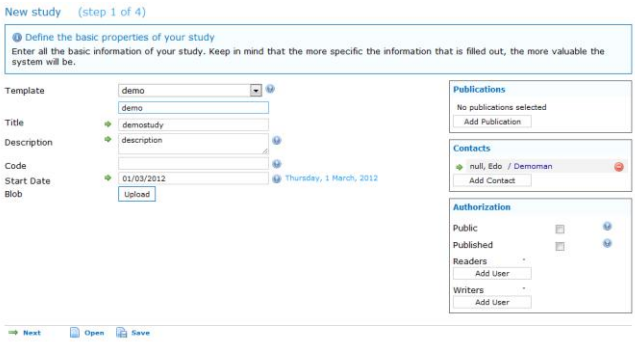
How to import my own samples and samplenames?

This chapter will show you how to import your own samples and sample names.

Start with Import > A complete study with straightforward design. This will start a short wizard.



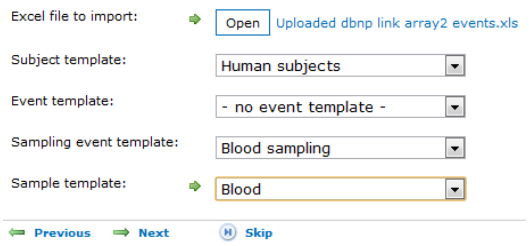
Define basic information of your study, and click next. Arrows indicate required fields:



Choose the excel file with appropriate subject and sample data, for example:

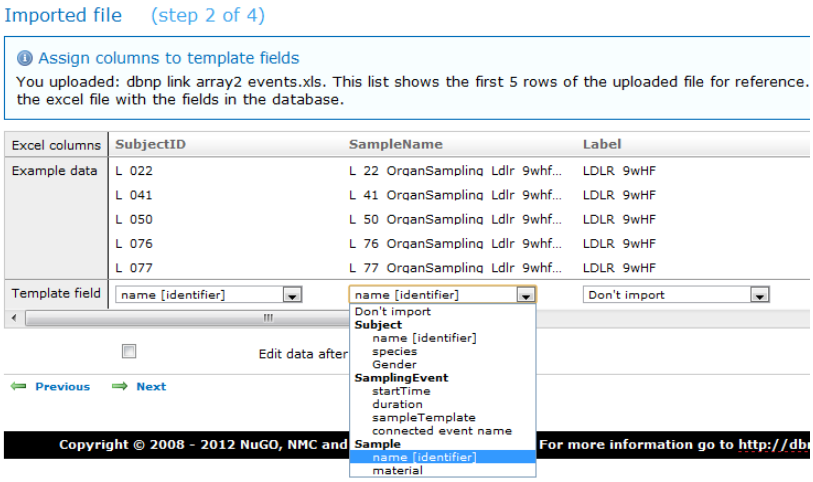
	A	B	C	D	E	F	G	H	I
1	SubjectID	SampleName	Label	Strain	genotype	Gender	name_old	starttime	endtim
2	L_022	L_22_OrganSampling_Ldlr_9whf_9wLiver	LDLR_9wHF	C57BL/6 N	LDLr-/-	Male	22	0w	0w
3	L_041	L_41_OrganSampling_Ldlr_9whf_9wLiver	LDLR_9wHF	C57BL/6 N	LDLr-/-	Male	41	0w	0w
4	L_050	L_50_OrganSampling_Ldlr_9whf_9wLiver	LDLR_9wHF	C57BL/6 N	LDLr-/-	Male	50	0w	0w
5	L_076	L_76_OrganSampling_Ldlr_9whf_9wLiver	LDLR_9wHF	C57BL/6 N	LDLr-/-	Male	76	0w	0w

Describe your subjects, samples and events, by using existing templates:



The important step during import is to map your columns with GSCF terms, to indicate how your excel sheet is structured. Any data you want to skip can be indicated by selecting 'Don't import'.

For example, choose subject name where your subject IDs and 'sample name [identifier]', where your samplenames are:



Fill in missing values, and select assays if you wish. The wizard will then give a study summary. After ‘Save’, you can view your study:

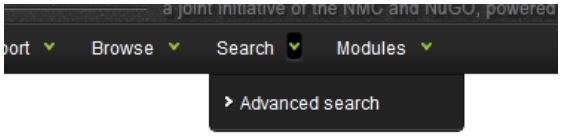
Show Study

Study Information | Subjects | Events timeline | Events table | Assays | Samples | Persons | Publications

Parent Subject	Parent Sampling Event	Parent Event Group	name
L_103	Blood sampling at 0 seconds		L_103_OrganSampling_Ldlr_9whf_9wLiver
L_120	Blood sampling at 0 seconds		L_120_OrganSampling_Ldlr_9whf_9wLiver
L_201	Blood sampling at 0 seconds		L_201_OrganSampling_Ldlr_16whf_16wLiver
L_202	Blood sampling at 0 seconds		L_202_OrganSampling_Ldlr_16whf_16wLiver
L_203	Blood sampling at 0 seconds		L_203_OrganSampling_Ldlr_16whf_16wLiver
L_204	Blood sampling at 0 seconds		L_204_OrganSampling_Ldlr_16whf_16wLiver
L_205	Blood sampling at 0 seconds		L_205_OrganSampling_Ldlr_16whf_16wLiver
L_206	Blood sampling at 0 seconds		L_206_OrganSampling_Ldlr_16whf_16wLiver
L_207	Blood sampling at 0 seconds		L_207_OrganSampling_Ldlr_16whf_16wLiver
L_208	Blood sampling at 0 seconds		L_208_OrganSampling_Ldlr_16whf_16wLiver
L_209	Blood sampling at 0 seconds		L_209_OrganSampling_Ldlr_16whf_16wLiver
L_210	Blood sampling at 0 seconds		L_210_OrganSampling_Ldlr_16whf_16wLiver

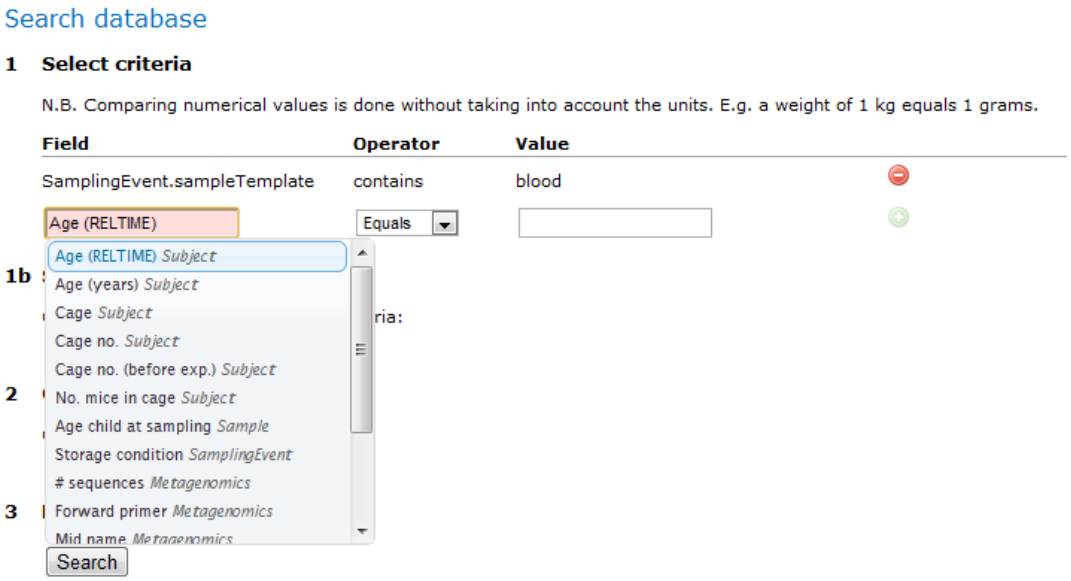
2. Advanced search

Data from GCSF and the assay modules can be queried by using the Advanced search. Use the menu on top of the page and click on Search > Advanced search.

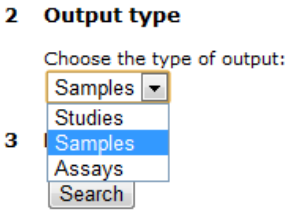


On this page you can search for *any of the parameters/fields used in GSCF and the modules*. This can be Age, Subject Name, Cholesterol (or other feature names used in Simple Assay module), or Classification (or other fieldnames used in Metagenomics module).

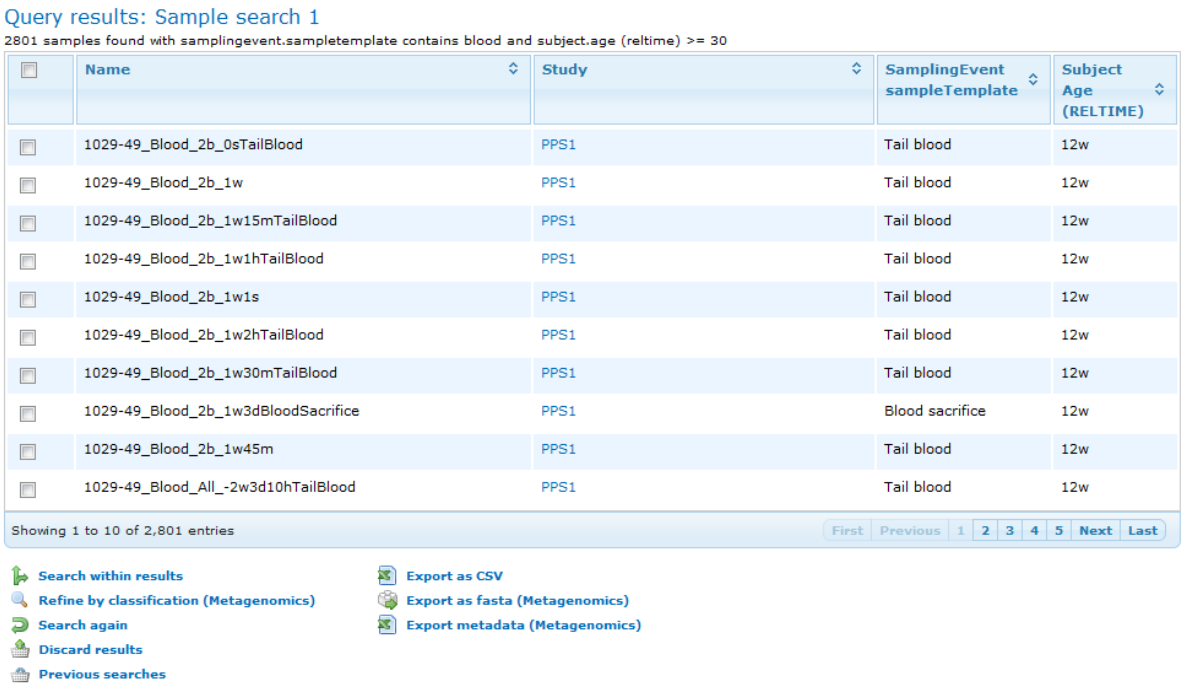
Search example: Let's assume you want to have a list of all blood samples, where subjects are of age 30 or higher (among all your studies). First, you have to specify parameters. Note that while you type for a fieldname, you will get a list of possible fieldnames that exist in the database:



Choose your output format: results can be shown on either study-, sample or assay level.



Click Search, and you will get a list of all the samples that match to the specified criteria. Results of this search can be exported, for example as CSV file.



Query results: Sample search 1

2801 samples found with samplingevent.sampletemplate contains blood and subject.age (reltime) >= 30

Name	Study	SamplingEvent sampleTemplate	Subject Age (RELTIME)
1029-49_Blood_2b_0sTailBlood	PPS1	Tail blood	12w
1029-49_Blood_2b_1w	PPS1	Tail blood	12w
1029-49_Blood_2b_1w15mTailBlood	PPS1	Tail blood	12w
1029-49_Blood_2b_1w1hTailBlood	PPS1	Tail blood	12w
1029-49_Blood_2b_1w1s	PPS1	Tail blood	12w
1029-49_Blood_2b_1w2hTailBlood	PPS1	Tail blood	12w
1029-49_Blood_2b_1w30mTailBlood	PPS1	Tail blood	12w
1029-49_Blood_2b_1w3dBloodSacrifice	PPS1	Blood sacrifice	12w
1029-49_Blood_2b_1w45m	PPS1	Tail blood	12w
1029-49_Blood_All_-2w3d10hTailBlood	PPS1	Tail blood	12w

Showing 1 to 10 of 2,801 entries

First Previous 1 2 3 4 5 Next Last

Search within results
Refine by classification (Metagenomics)
Search again
Discard results
Previous searches

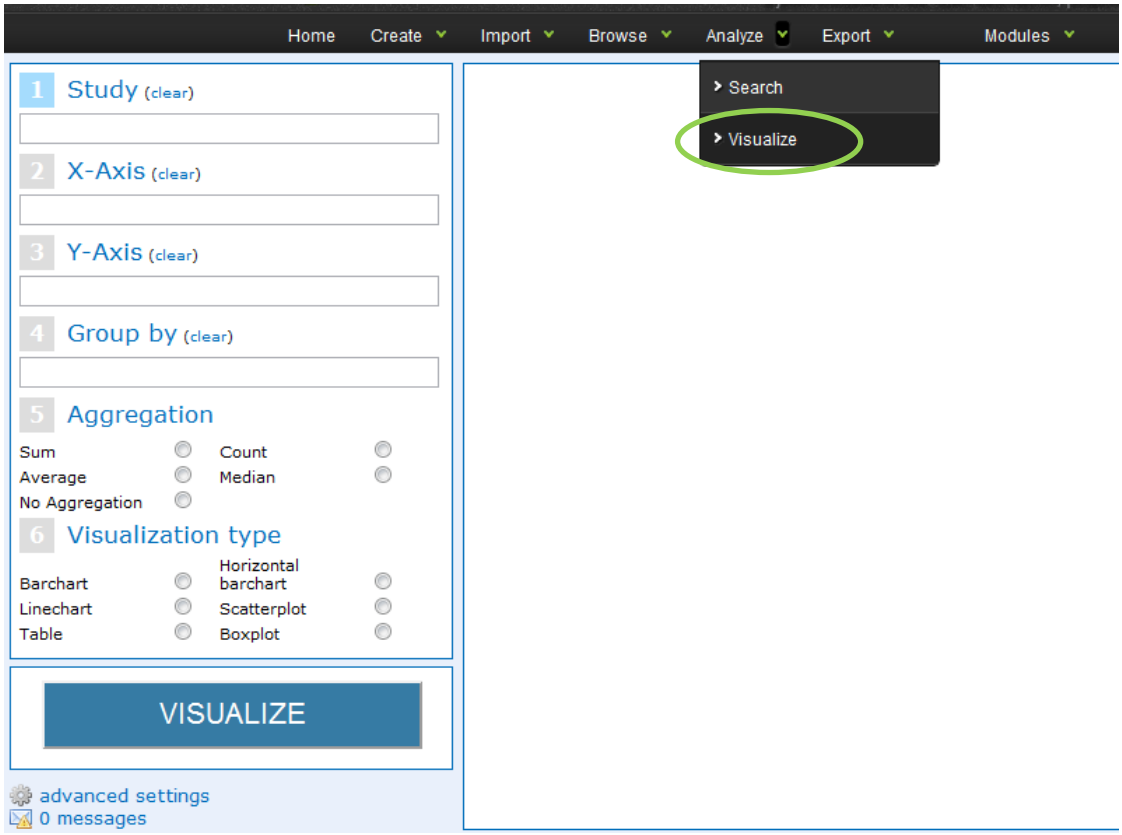
Export as CSV
Export as fasta (Metagenomics)
Export metadata (Metagenomics)

- Note: especially when you choose *sample* as the output format, this search can take several minutes, since the number of samples can be very high.

3. Visualize

The visualization tool GSCF allows you to visualize any data that is entered into the database. This can be subject-related data (such as weight, BMI or average age), but also SAM data (such as average leptin-levels measured, grouped by event group).

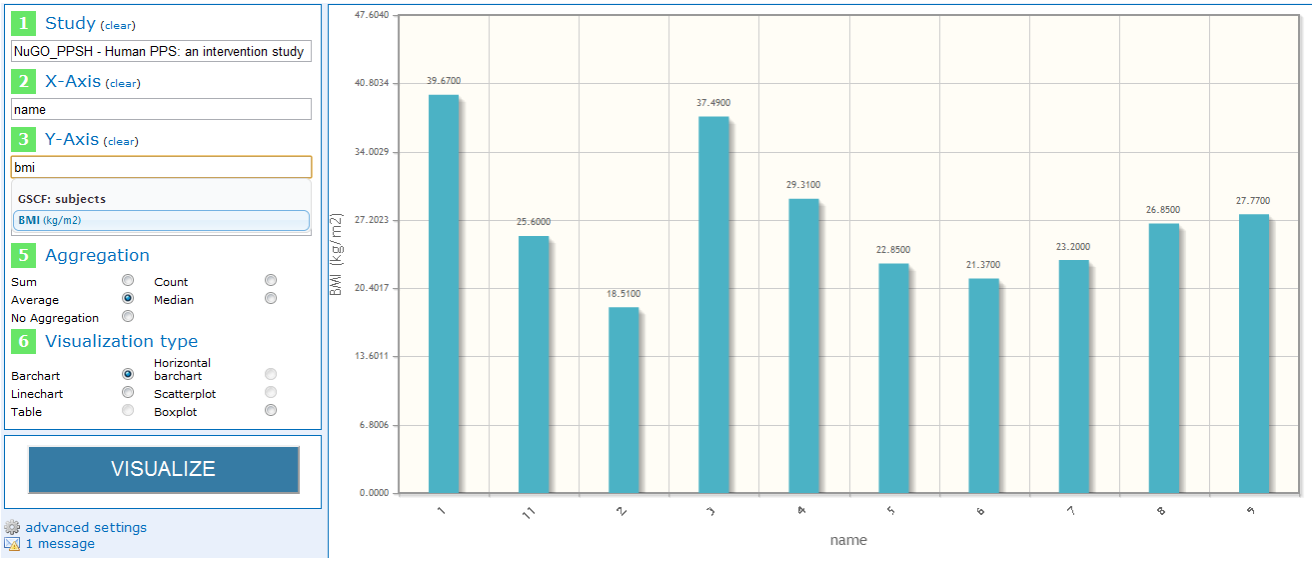
To enter the visualization tool, go to Analyze > Visualize, and you will see the control panel for visualizations on the left:



3.1 How to get GSCF data visualized?

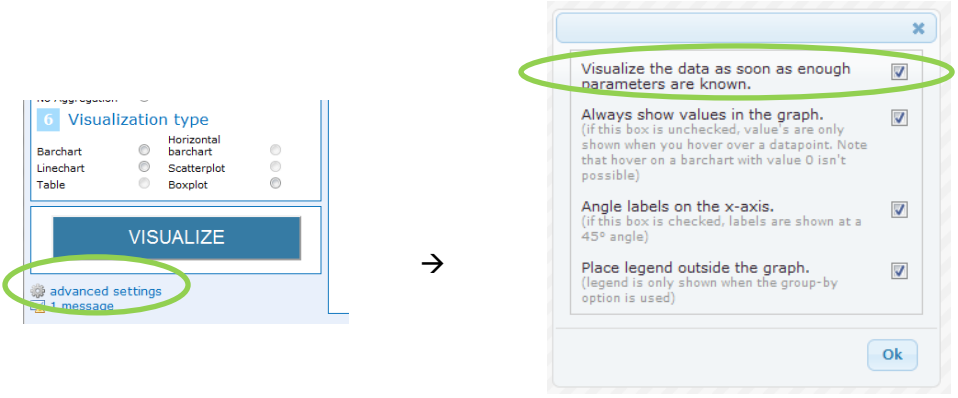
First, select your study and choose parameters for the X- and Y-axis. You can use the search bar and just type the field you want to use. The Group by functionality is optional, so you can leave this unchanged.

Step 5 will allow you to specify how to aggregate the data, but average is selected by default. If there are multiple visualization types possible, you have to indicate how you want to view it. If all parameters are set, the graph will automatically be generated.



Below the 'Visualize' button there are links for advanced settings and messages. If your data is invalid to be visualized, or cannot be found, a message will appear. Click on the link to read the full message.

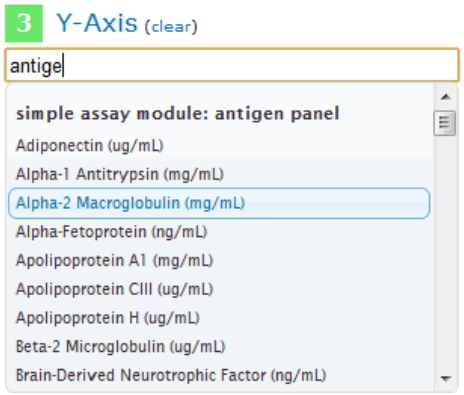
Click on advanced settings for more interface options. Especially with studies that have a lot of data, the first item is important: by default the tool will show a graph as soon as enough parameters is known. To visualize a graph manually deselect this item:



3.2 How to get module data visualized?

Instead of GSCF fields, you can also select fields from the modules that you have entered data into. For instance, you can select a feature from SAM (simple assay module), by selecting this on the Y-axis. All features are grouped together by assay name.

You can also type the assay name to see what features are possible:



Aggregation

Several aggregation options are possible. By default 'Average' is selected, however you can choose any of the options that are enabled. The graph will reload when you choose another type of calculation of your data.

Visualization types

The tool enables you to visualize your data in different kinds. Depending on the type of data you selected, the types are functional. For example, if you choose to have categorical data on the X-axis (e.g. event group name, or subject number) you cannot choose for a scatterplot.

VISUALIZATION TYPES

X-AXIS	Y-AXIS	TYPE
Numerical	Numerical	Scatterplot, Linechart
Numerical	Categorical	Horizontal Barchart
Categorical	Numerical	Barchart, Linechart
Categorical	Categorical	Table

4. Export data

Both GSCF as well as Assay data can be exported.

4.1 Export GSCF data

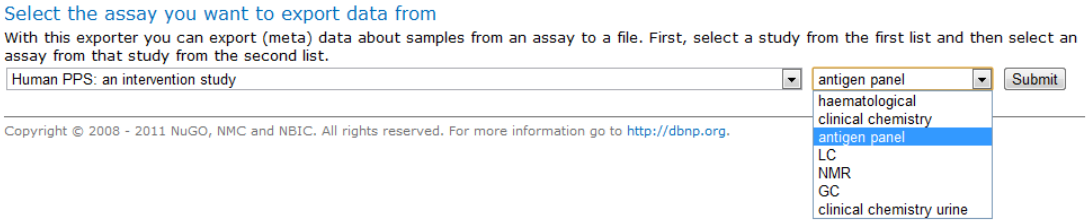
A whole study with GSCF data can be exported as a SimpleTox Excel file. To do this, navigate to Export > Export studies as SimpleTox Excel file. Select which studies you want to export, and click Export. The output will look like this:

	A	B	C	D	E	F	G	H	I	J	K	L
1	SubjectID	DataFile	HybName	SampleNa	ArrayType	Label	StudyTitle	Array_ID	Species	name	species	Body Weig C
2	Subject 6			Subject6_WholeBody		control	PPS3		Mus musc	Subject 6	Mus musculus	
3	Subject 8			Subject8_WholeBody		control	PPS3		Mus musc	Subject 8	Mus musculus	
4	Subject 5			Subject5_WholeBody		control	PPS3		Mus musc	Subject 5	Mus musculus	
5	Subject 7			Subject7_WholeBody		control	PPS3		Mus musc	Subject 7	Mus musculus	
6	Subject 10			Subject10_WholeBod		control	PPS3		Mus musc	Subject 10	Mus musculus	
7	Subject 3			Subject3_WholeBody		Minocyclin	PPS3		Mus musc	Subject 3	Mus musculus	
8	Subject 9			Subject9_WholeBody		Minocyclin	PPS3		Mus musc	Subject 9	Mus musculus	
9	Subject 1			Subject1_WholeBody		control	PPS3		Mus musc	Subject 1	Mus musculus	
10	Subject 4			Subject4_WholeBody		Minocyclin	PPS3		Mus musc	Subject 4	Mus musculus	
11	Subject 2			Subject2_WholeBody		control	PPS3		Mus musc	Subject 2	Mus musculus	
12	Subject 16			Subject16_WholeBod		HF-AL	PPS3		Mus musc	Subject 16	Mus musculus	
13	Subject 12			Subject12_WholeBod		cHF	PPS3		Mus musc	Subject 12	Mus musculus	
14	Subject 17			Subject17_WholeBod		cHF	PPS3		Mus musc	Subject 17	Mus musculus	
15	Subject 14			Subject14_WholeBod		HF-AL	PPS3		Mus musc	Subject 14	Mus musculus	
16	Subject 13			Subject13_WholeBod		HF-AL	PPS3		Mus musc	Subject 13	Mus musculus	
17	Subject 20			Subject20_WholeBod		Chow	PPS3		Mus musc	Subject 20	Mus musculus	

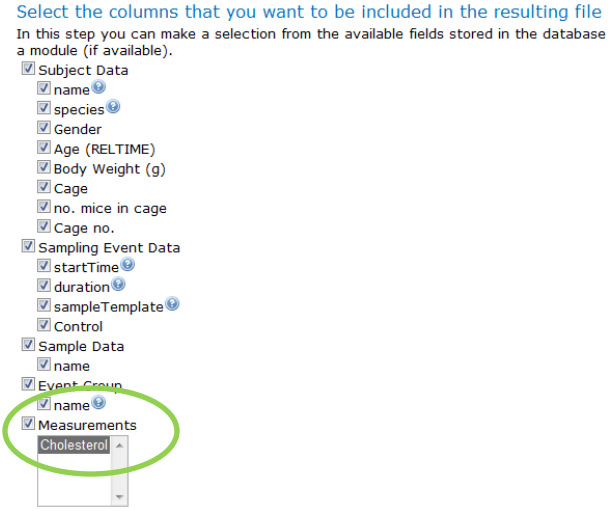
4.2 Export Assay data

If you want to export a study, but also the data from the assays that belong to this study, use Export > Export Assay data to file. This will allow you to export measurements from SAM, for example.

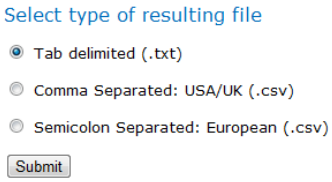
Select the study and assay for export and Submit.



In the next screen you can specify in more detail which fields you want to export. If you want to export SAM Assay data, make sure that Measurements is checked:



Finally, select the file type. All types can be opened in Excel.



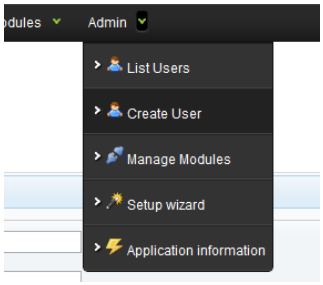
5. Access control

5.1 Admin and users

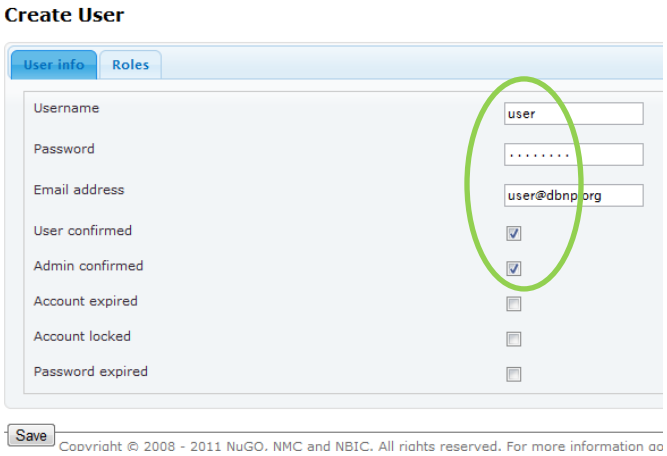
GSCF makes a distinction between administrators and users. When a new users registers, he or she will automatically become a normal user. When a new user is added by the administrator, he can give this user either normal user rights, or an administrator role.

How to create a new user or admin?

Go to Admin > Create User

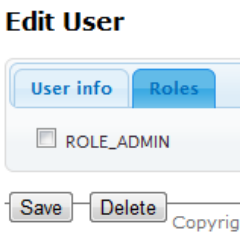


All three, username, password and email address are required fields. To finish this user, make sure User confirmed and Admin confirmed are checked, and click Save.



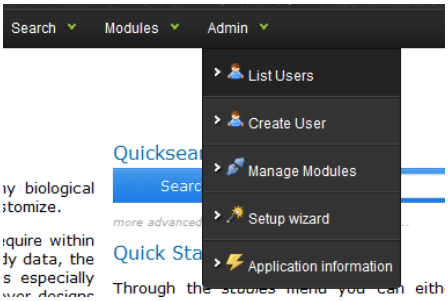
How to give a user administrator rights?

Only after a user has been created, the option of giving administrator rights is enabled. This can be found in the Roles tab, while editing a user:

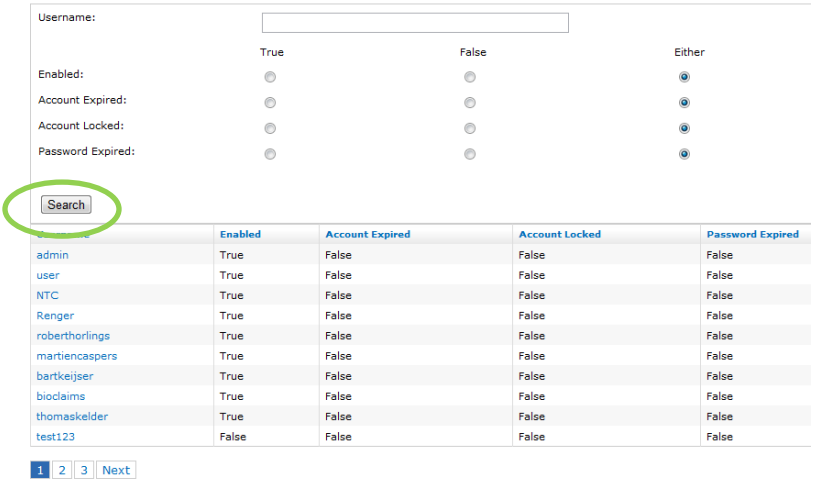


How to see all users and their access rights?

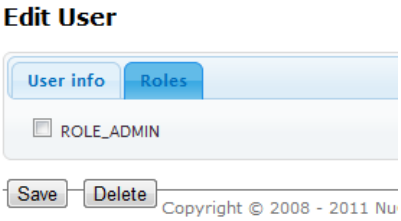
Go to Admin > List Users



You will enter a search page, but if you search without any given character for username, you get a full list of all known users and administrators:



To see the role for a certain user name, click on that name, and go to the tab Roles. Here you can see if this user has admin rights.



5.2 Readers and writers

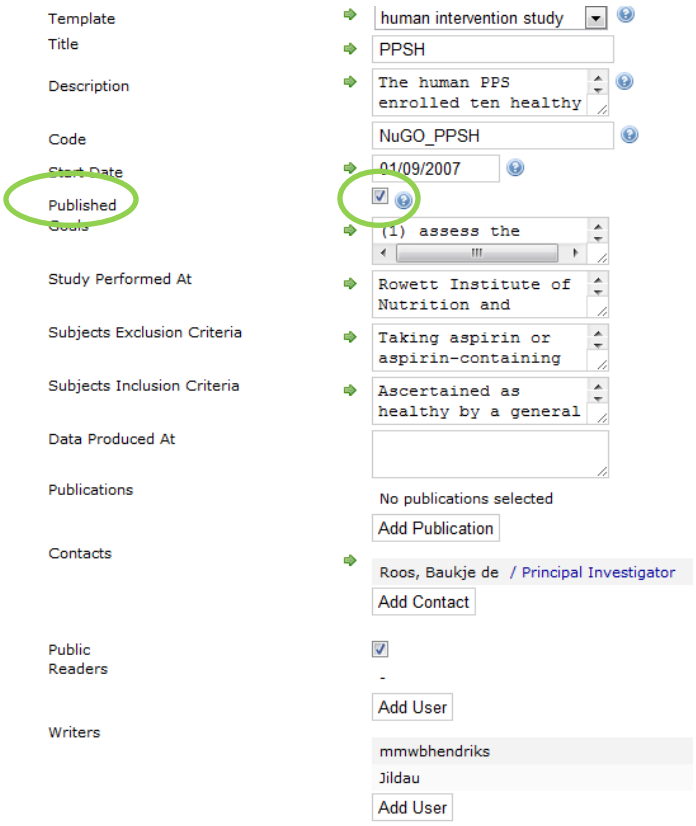
When a new study is created, other users can become Reader or Writer of that study. Readers are able to see the study, but cannot edit it. Writers can also edit the study. Before you assign Readers and Writers, make sure your study is published first.

How to publish a created study?

Before a study is actual visible to study readers and writers, make sure that the box Published is checked. This can be done in the first page of the study create Wizard.

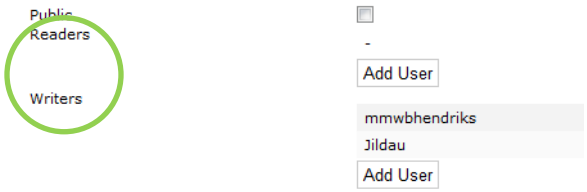
This determines whether this study is published (accessible for the study readers and, if the study is public, for anonymous users). A study can only be published if it meets certain quality criteria, which will be checked upon save.

The option is maybe a bit hidden and easily skipped, but can be found here:

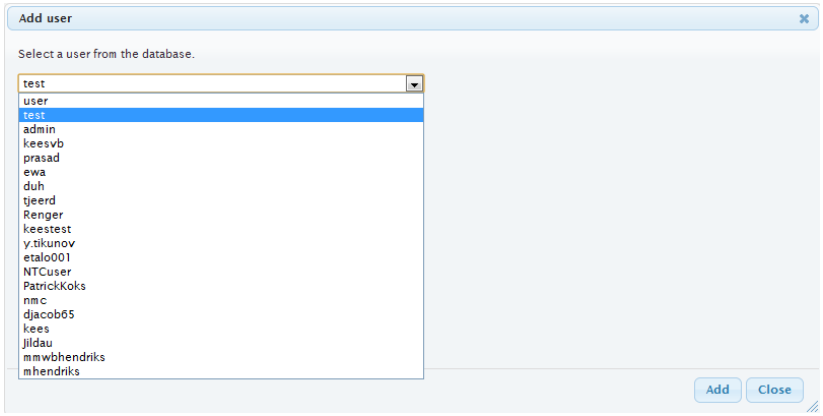


How to add users as readers or writer of a study?

In the first step of the study create wizard, you can assign users as public Readers or as Writers of the study.

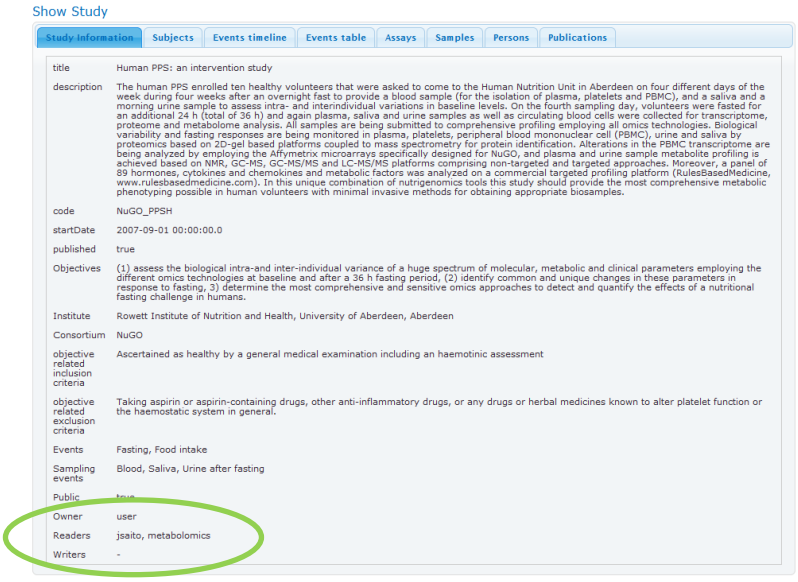


Click on Add User, and select the user you want to become a Reader, and click on Add to add this person as Reader. Follow the same procedure for Writers, to assign a user as Writer.



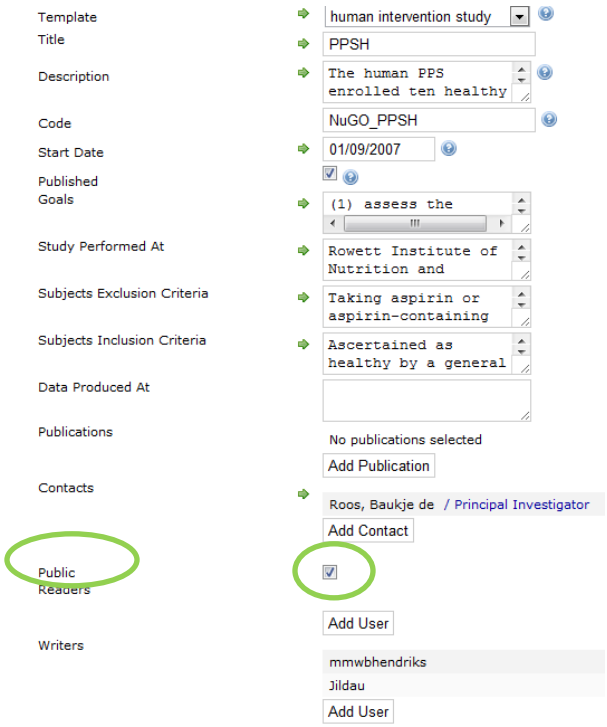
How to see Readers and Writers of a study?

To see who are Readers or Writers of a study, go to the study list (Browse > All Studies) and view a study. On the first tab, Study Information, you can see a list of the study owner, and the assigned Readers and Writers:



How to make a study publicly available?

If you want to make a study publicly readable, so even for non-DSP users, you can indicate this also in the first step. Make sure that Public is checked:

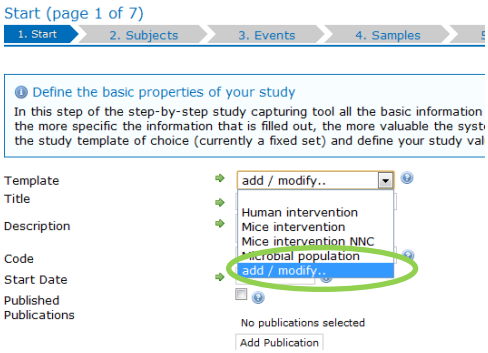


6. Templates

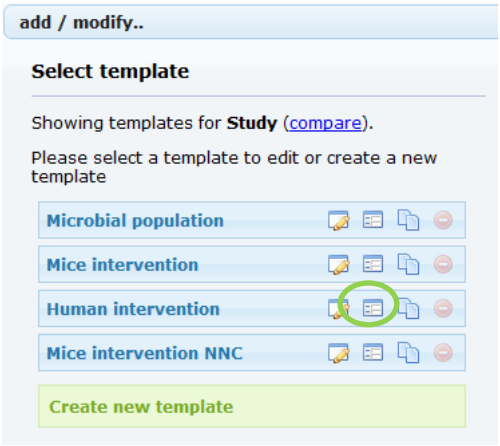
In order to allow flexibility to capture all information you require within a study, and to make it possible to compare studies or study data, the system uses customizable templates and ontologies. Templates are available at all steps of the create study wizard.

For example, to indicate the type of study design, choose an existing Study template, or 'add/modify templates'. This will show a popup where you can define the fields you want to appear next to default fields for study details:

Choose Add/Modify to add or edit templates:

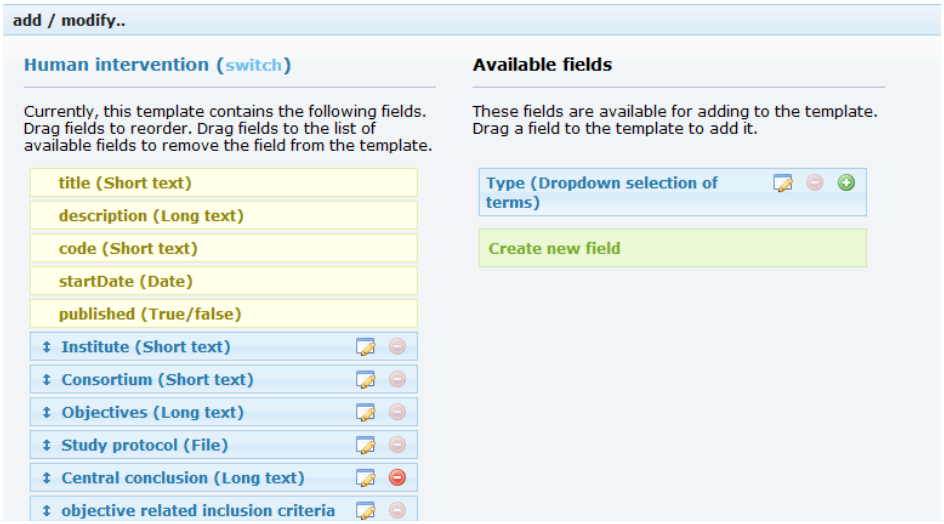


Now you can see a list of existing templates:



By clicking the first icon you can change the name and description. For each template you can also describe additional fields. Click on the second icon to modify template fields.

In this screen you can create new fields, indicate what is required for these fields, and you can organize all fields:



Other parts within the GSCF study design that apply the template system are: Subjects (Species template), Events, Sampling Events, Samples and Assays.

FAQs

Where can I see a list of my studies?

Click on Browse > My studies to view your studies. It is possible that other users indicate you are reader of a study. These studies also appear in this list.

Where can I update study information?

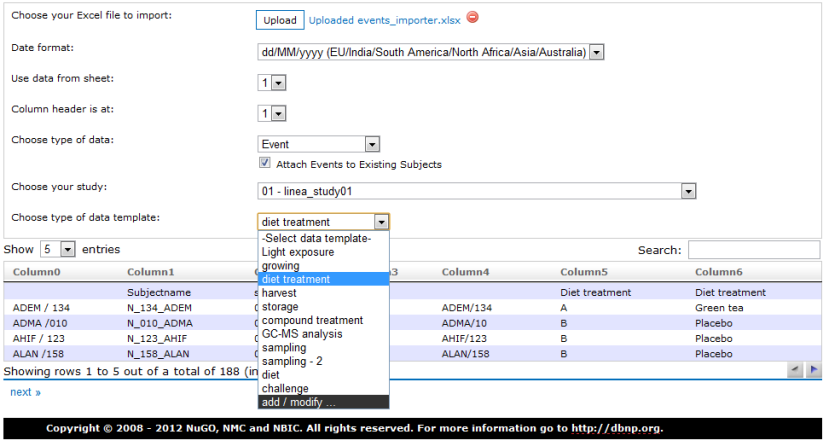
Click on Browse > My studies, and use the pencil on the left to edit a certain study. Just as creating a new study, this will start the study wizard. Using 'next' and 'previous' you can navigate through your study and update any information from this study.

How can I import events, or sampling events?

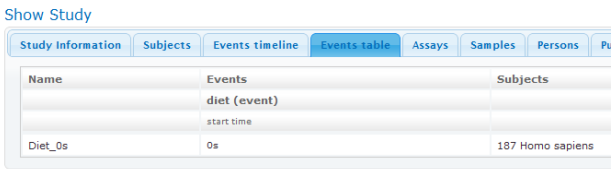
Events or sampling events can be imported using a short wizard (the same way you can import subjects or samples, described in chapter 1.3). Use the following format if you want to import events. Describe starttime, endtime, and a column for event description (second Diet treatment).

Subjectname	starttime	endtime		Diet treatment	Diet treatment
N_134_ADEM	0s	12w	ADEM/134	A	Green tea
N_010_ADMA	0s	12w	ADMA/10	B	Placebo
N_123_AHIF	0s	12w	AHIF/123	B	Placebo
N_158_ALAN	0s	12w	ALAN/158	B	Placebo

Use Import> Import part of study design, choose your file and indicate where your data begins. If your excel sheet contains subject names, you can link these events to the subjects you have in GSCF:



The result is one event, in which all subjects take part:



How can I add more samples to my study?

Click on Browse > My studies, and use the pencil on the left to edit a certain study. Samples are created as a combination of a sampling event, a subject and an event group. Change any of these to your needs and this will create more samples.

For example, if you miss a certain subject, add this subject in Step 2. Then navigate by clicking 'next' and this subject can be selected into a certain eventgroup. By clicking 'next' GSCF now creates new samples based on the new subject that was added to an eventgroup.

How can I add more samples to an assay?

See previous question how to add more samples into your study. Then go to Step 5, and click 'next'. Now you reach the in-between step, where you can add samples to an assay.

How are sample names created in GSCF?

Each sample name created in GSCF is composed of several fields. For example, the name 11_UrineAfterFasting_AllPersons_-1d has been created from:

Subject name: 11

Event group: All Persons

Sampling event: Urine After Fasting

Starttime Sampling event: -1 day

If there are more samples with the same content, then sample names get the addition _1, _2,_3.

Can I change the sample names created by GSCF?

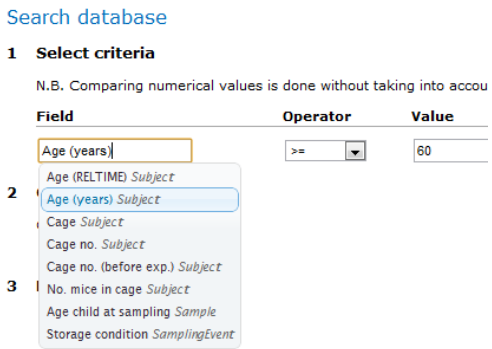
Yes, you can change the names that are automatically created by GSCF. However, especially when you have a long list of samples, this has to be done all manually. Go to the Study Edit > Samples.

Another option is to import samples, and thereby use your own samplenames. During the import you can attach samples to corresponding subjects that you have in GSCF. See chapter 1, page 13, for more details on sample import.

Where can I search among multiple studies?

Click on Search > Advanced Search. Select your GSCF fields and values you want to search for. Then select type of output (based on Studies, Samples or Assays).

For example, show me all samples from subjects that are at least 62 years of age;



This will result in the following output:

Query results: Sample search 7

38 samples found with subject.age (years) >= 60

Name	Study	Subject Ag (years)
18-144_Blood_Placebo_15w	Relation between reduction of the inflammatory status and glucose metabolism in healthy overweight men (Diclofenac)	60
18-144_Blood_Placebo_-	Relation between reduction of the inflammatory status and glucose metabolism in healthy overweight men (Diclofenac)	60
18-144_Blood_Placebo_0s	Relation between reduction of the inflammatory status and glucose metabolism in healthy overweight men (Diclofenac)	60
18-144_Blood_Placebo_-_5	Relation between reduction of the inflammatory status and glucose metabolism in healthy overweight men (Diclofenac)	60
18-144_Blood_Placebo_-_4	Relation between reduction of the inflammatory status and glucose metabolism in healthy overweight men (Diclofenac)	60
18-144_Blood_Placebo_-_8	Relation between reduction of the inflammatory status and glucose metabolism in healthy overweight men (Diclofenac)	60

My samples do not appear in the module. What can I do?

For samples to appear in a module attached to GSCF, make sure you have selected the right module for this assay. For example, make sure you selected module Simple Assay Module when you trying to view samples in this module (SAM). Second, make sure the right samples are actually in this assay. You can check this by clicking ‘next’ in Step 5. Here you can see a list of which samples are in what assay.

If the right samples do not exist for this assay and module, add more samples: Samples are created as a combination of a sampling event, a subject and an event group. Change any of these to your needs and this will create more samples.

> For example, if you miss a certain subject, add this subject in Step 2. Then navigate by clicking ‘next’ and this subject can be selected into a certain eventgroup. By clicking ‘next’ GSCF now creates new samples based on the new subject that was added to an eventgroup.